Assignment 5 - CSC/DSC 462 - Fall 2018 - Due December 11

In the following questions, do not use any continuity correction procedure unless instructed to do so in the question. The points awarded for each question are as follows: Q2, Q3, Q5, Q7 are worth 10 points; Q1, Q6 are worth 15 points; Q4 is worth 20 points.

Q1: Suppose a binomial random variable $X \sim bin(n, p)$ is observed to be $X = 150$, with sample size $n = 235$.

(a) Construct a level 0.95 confidence interval for p.

(b) Test hypothesis $H_o : p \leq 0.65$ against $H_a : p > 0.65$. Report a P-value. Is the null hypothesis rejected at a significance level of $\alpha = 0.05$? Do the test twice, without and with the continuity correction.

Q2: We are given an *iid* sample from a normal distribution $N(\mu, \sigma^2)$.

$$151.7, 146.3, 147.7, 152.9, 151.2, 148.0, 153.7, 151.0,$$

of sample size $n = 8$.

(a) Calculate a level $1 - \alpha$ upper confidence bound for σ.

(b) Suppose a sample of size n from the same normal distribution is being planned. What sample size is needed to be able to assert with a confidence level of 95% that the sample mean $\bar{X}$ will be within 0.5 units of the true mean with a probability of 0.99 (note that this is not quite the same as determining a sample size for a confidence interval of a given margin of error).

Q3: We are given independent samples of size $n_1 = 24$ and $n_2 = 47$ from two normally distributed populations. Suppose we observe sample variances $S_1^2 = 5.476$ and $S_2^2 = 39.942$. Do a hypothesis test of

$$H_o : \sigma_2^2 = \sigma_1^2$$
$$H_a : \sigma_2^2 \neq \sigma_1^2$$

using an $\alpha = 0.1$ significance level. Give explicitly the rejection regions, and also report a P-value.

Q4: Suppose the ability to reduce asthma symptoms of an experimental treatment is compared to a standard treatment. The experimental treatment is administered to $n_1 = 178$ subjects, and the standard treatment is administrated to $n_2 = 79$ subjects. For each subject, whether or not the treatment succeeded in reducing asthma symptoms was observed. Suppose the results are summarized in the following contingency table:

	Experimental Treatment	Standard Treatment	
Reduced asthma symptoms	154	53	207
Asthma symptoms unaffected	24	26	50
Total	178	79	257

(a) Test hypothesis $H_o : p_1 = p_2$ against $H_a : p_1 \neq p_2$, where p_1, p_2 are the respective proportion of subjects for which symptoms were reduced for each treatment. Use a two-sample difference in proportion test. Report a P-value. Is the null hypothesis rejected at a significance level of $\alpha = 0.05$?

(b) Construct a level 0.95 confidence interval for the log odds ratio of *Reduced asthma symptoms* between groups *Experimental Treatment* and *Standard Treatment*. Can you reject the null hypothesis H_o : $OR = 1$ against $H_a : OR \neq 1$ at significance level of $\alpha = 0.05$?

(c) Suppose we may interpret the table as a contingency table with $n_r = 2$ rows and $n_c = 2$ columns based on a random sample of size $n = 257$. Assume the probability of cell i, j are given by $p_{i,j}$. The population frequencies for the marginal row i and column j categories are given by r_i and c_j, respectively. Use a χ^2 test for the null hypothesis of row and column independence $H_o : p_{i,j} = r_i c_j$ for all i, j. Use significance level $\alpha = 0.05$.

(d) Verify that the null hypotheses of all three tests are the same.

Q5: According to the theory of Mendelian inheritance, the phenotypic ratios of a dihybrid cross (of which there are 4) are given by the ratios 9:3:3:1. Suppose in a sample of $n = 78$ species of plants believed to be a dihybrid cross, the four relevant traits were observed with the frequencies reported in the following table.

Traits	1	2	3	4	Totals
Observed counts O_i	19	13	19	27	78
Hypothetical frequencies p_i	1/16	3/16	3/16	9/16	1.00
Observed frequencies $\hat{p}_i$	0.24	0.17	0.24	0.35	1.00

Hypothetical population frequencies p_i are also given in the table. Use a χ^2 test for null hypothesis

$$H_o : p_i \text{ are the true population frequencies.}$$

Use significance level $\alpha = 0.05$. Use Yate's correction procedure.

Q6: We are given two paired samples of sample size $n = 9$. The data is summarized in the table below. Suppose $\tilde{\mu}_D$ is the population median of the paired differences $D = X - Y$.

(a) Perform a lower tailed signed rank test using hypotheses $H_o : \tilde{\mu}_D = 0$ against $H_a : \tilde{\mu}_D > 0$ with significance level $\alpha = 0.05$. Use both a normal approximation and the exact method.

(b) Perform a sign test for the same hypotheses, and report a P-value. Do you reach a different conclusion than that of Part (a)?

	Sample 1 (X)	Sample 2 (Y)	$D = X - Y$
1	20.1	16.9	3.2
2	16.7	20.3	-3.6
3	18.3	16.5	1.8
4	14.4	14.9	-0.5
5	16.6	16.0	0.6
6	15.2	14.1	1.1
7	15.7	15.0	0.7
8	13.1	12.9	0.2
9	11.7	11.7	0.0

Q7: We are given two independent samples of sample sizes $n_1 = 5$, $n_2 = 12$. The data is summarized in the table below. Suppose $\tilde{\mu_i}$ is the population median of sample i. Perform a two-sided rank sum test using hypotheses $H_o : \tilde{\mu_1} - \tilde{\mu_2} = 0$ against $H_a : \tilde{\mu_1} - \tilde{\mu_2} \neq 0$. Use significance level $\alpha = 0.01$.

	1	2	3	4	5	6	7	8	9	10	11	12	$\tilde{X}_i$
Sample 1	24.5	19.3	32.5	28.5	23.9								24.5
Sample 2	32.5	31.6	40.4	35.8	34.9	40.7	39.5	41.0	35.1	38.8	39.1	41.0	38.95